03-29-01

PCT

FORM PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		ATTORNEY'S DOCKET NUMBER **PF-0609 USN** U.S. APPLICATION NO. (If known, see 37 CFR 1.5) TO BE ASSIGNED 09/806276
INTERNATIONAL APPLICATION NO. **PCT/US99/22908**	INTERNATIONAL FILING DATE **01 October 1999**	PRIORITY DATE CLAIMED **02 October 1998**

TITLE OF INVENTION
BONE MARROW-DERIVED SERUM PROTEINS

APPLICANT(S) FOR DO/EO/US
INCYTE PHARMACEUTICALS, INC.; TANG, Y. Tom; CORLEY, Neil C.; GUEGLER, Karl J.; LU, Aina M. Lu

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is the **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This is an express request to promptly begin national examination procedures (35 U.S.C. 371 (f)).
4. ☐ The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☐ is attached hereto (required only if not communicated by the International Bureau)
   b. ☐ has been communicated by the International Bureau.
   c. ☒ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. ☐ are attached hereto (required only if not communicated by the International Bureau).
   b. ☐ have been communicated by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 to 16 below concern document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.27 and 3.31 is included.
13. ☐ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information:

1) Transmittal Letter (2 pp, in duplicate)
2) Return Postcard
3) Express Mail Label No.: **EL 856 112 870 US**

09806276 . 032901

JC03 Rec'd PCT/PTO 27 MAR 2001

U.S. APPLICATION NO. (if known, see 37 CFR 1.5) TO BE ASSIGNED: 09/806276	INTERNATIONAL APPLICATION NO.: PCT/US99/22908	ATTORNEY'S DOCKET NUMBER PF-0609 USN

17. ☒ The following fees are submitted:

BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5):

Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO.....$1000.00

☐International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO..$860.00

International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO.....$710.00

☒**International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4).....$690.00**

☐International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4).....$100.00

ENTER APPROPRIATE BASIC FEE AMOUNT =				$690.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				$	
CLAIMS	NUMBER FILED		NUMBER EXTRA	RATE	
Total Claims	20	=	0	X **$ 18.00**	$
Independent Claims	2	=	0	X **$ 80.00**	$
MULTIPLE DEPENDENT CLAIM(S) (if applicable)				+ **$270.00**	$
TOTAL OF ABOVE CALCULATIONS =				$690.00	
☐ Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.				$	
SUBTOTAL =				$690.00	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest clailmed priority date (37 CFR 1492(f)). +				$	
TOTAL NATIONAL FEE =				$690.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by the appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +				$	
TOTAL FEES ENCLOSED =				$690.00	
				Amount to be Refunded:	$
				Charged:	$

a. ☐ A check in the amount of $_____ to cover the above fees is enclosed.

b. ☒ Please charge my Deposit Account No. 09-0108 in the amount of $ 690.00 to cover the above fees.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 09-0108. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

INCYTE GENOMICS, INC.
3160 Porter Drive
Palo Alto, CA 94304

SIGNATURE

NAME: Diana Hamlet-Cox

REGISTRATION NUMBER: 33,302

DATE: 27 March 2001

WO 00/20588 PCT/US99/22908
09/806276

BONE MARROW-DERIVED SERUM PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of bone marrow-derived serum proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cancer, immune disorders, infections, and vascular disorders.

BACKGROUND OF THE INVENTION

Bone marrow is the site of blood cell formation, or hematopoiesis, from birth throughout adulthood. Blood cells are comprised of diverse cell types including red blood cells (erythrocytes) and white blood cells (leukocytes), all of which are derived from a common progenitor stem cell. During hematopoiesis, the stem cell is stimulated to proliferate and differentiate by specific growth factors called colony-stimulating factors. Blood cell maturation then proceeds through various stages, each stage characterized by further commitment of an immature blood cell to a specific, terminally differentiated state. In addition to hematopoietic cells, bone marrow also contains blood vessels, nerves, fatty tissue, and stromal cells. Stromal cells produce a supporting meshwork of collagen fibers and other extracellular matrix components which are important for promoting the growth and differentiation of hematopoietic cells. Deregulation of hematopoiesis can lead to neoplastic conditions such as leukemia or lymphoma, while insufficient hematopoiesis can lead to anemia or immunodeficiency.

A novel protein, MSE55 (marrow stromal/endothelial cell protein, 55 kilodaltons), has been identified from human stromal cells (Bahou, W.F. et al. (1992) J. Biol. Chem. 267:13986-13992). MSE55 is specifically expressed in stromal cells and in endothelial cells which line blood vessels. Furthermore, MSE55 is detected at relatively high levels in the serum, suggesting that stromal cells and/or endothelial cells secrete MSE55 into the circulation. MSE55 cDNA contains a long 5′ untranslated region of about 350 base pairs followed by a 1,173-base pair open reading frame that potentially encodes a 391-amino acid polypeptide. The observed molecular weight of 55 kilodaltons exceeds the predicted molecular weight of 42 kilodaltons, suggesting that MSE55 may undergo post-translational modifications such as glycosylation. Although MSE55 is secreted into the serum, the predicted amino acid sequence does not contain a signal peptide. Lack of a signal peptide, however, is also observed in other serum proteins such as plasminogen activator inhibitor 2 and ovalbumin. Other features of the predicted MSE55 sequence include a serine- and glycine-rich N-terminal region, an internal region of proline- and alanine-rich tandem repeats, and two putative metal-binding motifs. MSE55 is cross-reactive with antibodies against another

stromal cell protein, hemonectin, which plays an important role in white blood cell adhesion and maturation. Likewise, MSE55 may also play a similar role in white blood cell hematopoiesis.

B-lymphocytes are one of several types of differentiated white blood cells which play a critical role in the immune response to microbial infections. B-lymphocytes enter the bloodstream from the bone marrow and populate the spleen, lymph nodes, and other lymphoid organs. In these organs, B-cells encounter and react to foreign antigens by expressing and secreting antibodies into the circulation. Antibodies, or immunoglobulins, recognize and bind to antigens on the surface of blood-borne microbes. Antigen binding triggers an immune response which leads to the destruction of the microbe.

The prototypical antibody is a tetramer consisting of two identical heavy polypeptide chains (H-chains) and two identical light polypeptide chains (L-chains) interlinked by disulfide bonds. This arrangement confers the characteristic Y-shape to antibody molecules. Antibodies are classified based on their H-chain composition. The five antibody classes, IgA, IgD, IgE, IgG and IgM, are defined by the α, δ, ε, γ, and μ H-chain types, respectively. There are two types of L-chains, κ and λ, either of which may associate as a pair with any H-chain pair. IgG, the most common class of antibody found in the circulation, is tetrameric as described above, while the other classes of antibodies are generally variants or multimers of this basic structure.

H-chains and L-chains each contain an N-terminal variable region and a C-terminal constant region. The sequence of the constant region, which consists of about 110 amino acids in L-chains and about 330 or 440 amino acids in H-chains, is nearly identical among H- or L-chains of a particular class. However, the sequence of the variable region, which consists of about 110 amino acids, differs among H- or L-chains of a particular class. Within each H- and L-chain variable region are three hypervariable regions of extensive sequence diversity, each consisting of about 5 to 10 amino acids. In the antibody molecule, the H- and L-chain hypervariable regions come together to form the antigen binding site. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing, New York NY, pp. 1206-1213, 1216-1217.)

Recombinant DNA technology has enabled the production of antibodies engineered for use as therapeutic and diagnostic agents. For example, rodent antibodies directed against human disease-associated proteins can be "humanized" by replacing their constant regions with those from human antibodies (Junghans, R.P. et al. (1990) Cancer Res. 50:1495-1502). The variable regions of these humanized antibodies recognize human disease-associated proteins, while the constant regions activate downstream effectors and prevent the antibodies themselves from being recognized as foreign in a human host. Humanized antibodies have proved to be effective therapeutic agents for the prevention of transplant rejection in primate model systems and for their

anti-proliferative activity in breast tumor cell lines (Brown, P.S. et al. (1991) Proc. Natl. Acad. Sci. USA 88:2663-2667).

The discovery of new bone marrow-derived serum proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cancer, immune disorders, infections, and vascular disorders.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, bone marrow-derived serum proteins, referred to collectively as "BMDSP" and individually as "BMDSP-1" and "BMDSP-2." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and fragments thereof.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-2, and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, and fragments

thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID 3-4, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and fragments thereof. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-2, and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of BMDSP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of BMDSP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and fragments thereof.

BRIEF DESCRIPTION OF THE FIGURES AND TABLE

Figures 1A, 1B, and 1C show the amino acid sequence (SEQ ID NO:1) and nucleic acid sequence (SEQ ID NO:3) of BMDSP-1. The alignment was produced using MACDNASIS PRO

software (Hitachi Software Engineering, S. San Francisco CA).

Figures 2A, 2B, 2C, 2D, 2E, and 2F show the amino acid sequence (SEQ ID NO:2) and nucleic acid sequence (SEQ ID NO:4) of BMDSP-2.

Figures 3A and 3B show the amino acid sequence alignment between BMDSP-2 (1859631; SEQ ID NO:2) and amino acids 1 through 218 of MSE55 (GI 338033; SEQ ID NO:5), produced using the multisequence alignment program of LASERGENE software (DNASTAR, Madison WI).

Table 1 shows the tools, programs, and algorithms used to analyze BMDSP, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"BMDSP" refers to the amino acid sequences of substantially purified BMDSP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic,

semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which, when bound to BMDSP, increases or prolongs the duration of the effect of BMDSP. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of BMDSP.

An "allelic variant" is an alternative form of the gene encoding BMDSP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding BMDSP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide the same as BMDSP or a polypeptide with at least one functional characteristic of BMDSP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding BMDSP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding BMDSP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent BMDSP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of BMDSP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine.

The terms "amino acid" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of BMDSP which are preferably at least 5 to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of BMDSP. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are

not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which, when bound to BMDSP, decreases the amount or the duration of the effect of the biological or immunological activity of BMDSP. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of BMDSP.

The term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind BMDSP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

The term "biologically active," refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic BMDSP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" or a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding BMDSP or fragments of BMDSP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the GELVIEW Fragment Assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

The term "correlates with expression of a polynucleotide" indicates that the detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding BMDSP, by northern analysis is indicative of the presence of nucleic acids encoding BMDSP in a sample, and thereby correlates with expression of the transcript from the polynucleotide encoding BMDSP.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the

natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" or "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR) which creates alignments between two or more sequences according to methods selected by the user, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) Gene 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may

contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" or "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" or "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of BMDSP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of BMDSP.

The phrases "nucleic acid" or "nucleic acid sequence," as used herein, refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, "fragments" refers to those nucleic acid sequences which comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:3-4, for example, as distinct from any other sequence in the same genome. For example, a fragment of SEQ ID NO:3-4 is useful in hybridization and amplification technologies

and in analogous methods that distinguish SEQ ID NO:3-4 from related polynucleotide sequences. A fragment of SEQ ID NO:3-4 is at least about 15-20 nucleotides in length. The precise length of the fragment of SEQ ID NO:3-4 and the region of SEQ ID NO:3-4 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment. In some cases, a fragment, when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-binding site, of the full-length polypeptide.

The terms "operably associated" or "operably linked" refer to functionally related nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization assay or microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding BMDSP, or fragments thereof, or BMDSP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" or "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "stringent conditions" refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent, e.g., formamide, temperature, and other conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of BMDSP polypeptides refers to an amino acid sequence that is altered by one or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g., replacement of glycine with tryptophan). Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

The term "variant," when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to BMDSP. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

THE INVENTION

The invention is based on the discovery of new human bone marrow-derived serum proteins (BMDSP), the polynucleotides encoding BMDSP, and the use of these compositions for the diagnosis, treatment, or prevention of cancer, immune disorders, infections, and vascular disorders.

Nucleic acids encoding the BMDSP-1 of the present invention were identified in Incyte Clone 135698H1 from the bone marrow cDNA library (BMARNOT02) using a computer search for nucleotide and/or amino acid sequence alignments. A consensus sequence, SEQ ID NO:3, was derived from the following overlapping and/or extended nucleic acid sequences: Incyte Clones 135698H1 (BMARNOT02), 1320039H1 (BLADNOT04), 792424T1 (PROSTUT03), 3430675T6 (SKINNOT04), and 2056224X14R1 (BEPINOT01).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:1, as shown in Figures 1A, 1B, and 1C. BMDSP-1 is 234 amino acids in length and has eight potential casein kinase II phosphorylation sites at T18, S34, S87, S96, T122, S182, T184, and S202 and two potential protein kinase C phosphorylation sites at S42 and S72. PFAM analysis indicates that BMDSP-1 contains two immunoglobulin domains from G36 to Q110 and from S147 to V216. Within and overlapping the latter domain are four immunoglobulin signatures as indicated by BLOCKS, MOTIFS, and PROFILESCAN analyses. These signatures include amino acid residues from D190 to E233, from S151 to A173, and from Y212 to F229. Likewise, BLAST searches of protein databases indicate that BMDSP-1 has chemical and structural similarity with immunoglobulin κ light chain. A fragment of SEQ ID NO:3 from about

nucleotide 339 to about nucleotide 407 is useful in hybridization or amplification technologies to identify SEQ ID NO:3 and to distinguish between SEQ ID NO:3 and a related sequence. Northern analysis shows the expression of this sequence in various libraries, at least 63% of which are associated with cancer or cell proliferation and at least 37% of which are associated with inflammation or trauma. In particular, 26% of the libraries expressing BMDSP-1 are derived from gastrointestinal tissue, and 26% are derived from reproductive tissue.

Nucleic acids encoding the BMDSP-2 of the present invention were identified in Incyte Clone 1859631H1 from the prostate cDNA library (PROSNOT18) using a computer search for nucleotide and/or amino acid sequence alignments. A consensus sequence, SEQ ID NO:4, was derived from the following overlapping and/or extended nucleic acid sequences: Incyte Clones 1859631H1 (PROSNOT18), 1426610F1 (SINTBST01), 1511409F1 (LUNGNOT14), 1349506F1 (LATRTUT02), 1005112H1 (BRSTNOT03), 1980941T6 (LUNGTUT03), 1213785R1 (BRSTTUT01), 1544433R1 (PROSTUT04), and 1440896F1 (THYRNOT03).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:2, as shown in Figures 2A, 2B, 2C, 2D, 2E, and 2F. BMDSP-2 is 254 amino acids in length and has two potential N-glycosylation sites at N88 and N162 and eight potential casein kinase II phosphorylation sites at S89, S130, S158, S191, T209, T218, S223, and S241. As shown in Figures 3A and 3B, BMDSP-2 has chemical and structural similarity with MSE55. BMDSP-2 shares 25% amino acid sequence identity with the N-terminal region of MSE55 from amino acids 1 through 218 (GI 338033; SEQ ID NO:5). In particular, the region of BMDSP-2 from L26 to G68 shares 60% identity with the homologous region of MSE55 and contains six glycine residues, consistent with the glycine-rich nature of the MSE55 N-terminal region. Furthermore, nine out of eleven residues which comprise the second putative metal binding motif in MSE55 are conserved in BMDSP-2 from Q235 to E245.

Like the cDNA encoding MSE55, the cDNA encoding BMDSP-2 (SEQ ID NO:4) contains a long 5′ untranslated region of about 415 base pairs. A fragment of SEQ ID NO:4 from about nucleotide 383 to about nucleotide 442 is useful in hybridization or amplification technologies to identify SEQ ID NO:4 and to distinguish between SEQ ID NO:4 and a related sequence. Northern analysis shows the expression of this sequence in various libraries, at least 67% of which are associated with cancer or cell proliferation and at least 34% of which are associated with inflammation or trauma. In particular, 27% of the libraries expressing BMDSP-2 are derived from reproductive tissue, and 20% are derived from cardiovascular tissue.

The invention also encompasses BMDSP variants. A preferred BMDSP variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least

about 95% amino acid sequence identity to the BMDSP amino acid sequence, and which contains at least one functional or structural characteristic of BMDSP.

The invention also encompasses polynucleotides which encode BMDSP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:3-4, which encodes BMDSP.

The invention also encompasses a variant of a polynucleotide sequence encoding BMDSP. In particular, such a variant polynucleotide sequence will have at least about 80%, more preferably at least about 90%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding BMDSP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:3-4 which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:3-4. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of BMDSP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding BMDSP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring BMDSP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode BMDSP and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring BMDSP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding BMDSP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding BMDSP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode BMDSP and BMDSP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding BMDSP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:3 and SEQ ID NO:4 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3

mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 (Hamilton, Reno NV), Peltier Thermal Cycler 200 (PTC200; MJ Research, Watertown MA) and the ABI CATALYST 800 (Perkin-Elmer). Sequencing is then carried out using either ABI 373 or 377 DNA sequencing systems (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding BMDSP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic

DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode BMDSP may be cloned in recombinant DNA molecules that direct expression of BMDSP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express BMDSP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter BMDSP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding BMDSP may be synthesized, in whole or in

part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, BMDSP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of BMDSP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g, Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active BMDSP, the nucleotide sequences encoding BMDSP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding BMDSP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding BMDSP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding BMDSP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding BMDSP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989)

Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding BMDSP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding BMDSP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding BMDSP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding BMDSP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of BMDSP are needed, e.g. for the production of antibodies, vectors which direct high level expression of BMDSP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of BMDSP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of BMDSP. Transcription of sequences encoding BMDSP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987)

EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding BMDSP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses BMDSP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of BMDSP in cell lines is preferred. For example, sequences encoding BMDSP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*$^-$ or *apr*$^-$ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides, neomycin and G-418;

and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding BMDSP is inserted within a marker gene sequence, transformed cells containing sequences encoding BMDSP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding BMDSP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding BMDSP and that express BMDSP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of BMDSP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on BMDSP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art

and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding BMDSP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding BMDSP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding BMDSP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode BMDSP may be designed to contain signal sequences which direct secretion of BMDSP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding BMDSP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric BMDSP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of BMDSP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione

S-transferase (GST), maltose binding protein (MBP), thioredoxin (T.x), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the BMDSP encoding sequence and the heterologous protein sequence, so that BMDSP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled BMDSP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract systems (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably ^{35}S-methionine.

Fragments of BMDSP may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Various fragments of BMDSP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of BMDSP-1 and immunoglobulin κ light chain. The expression of BMDSP-1 is closely associated with cancer and immune disorders. Therefore, BMDSP-1 appears to play a role in cancer, immune disorders, and infections. Furthermore, chemical and structural similarity exists between regions of BMDSP-2 and MSE55. BMDSP-2 is expressed in cardiovascular tissue and tissues associated with cancer and immune disorders. Therefore, BMDSP-2 appears to play a role in cancer, immune disorders, and vascular disorders. In the treatment of cancer, immune disorders, infections, and vascular disorders associated with increased BMDSP expression or activity, it is desirable to decrease the expression or activity of BMDSP. In the treatment of cancer, immune disorders, infections, and vascular disorders associated with decreased BMDSP

expression or activity, it is desirable to increase the expression or activity of BMDSP.

Therefore, in one embodiment, BMDSP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of BMDSP. Examples of such disorders include, but are not limited to, cancers such as adenocarcinoma, melanoma, sarcoma, teratocarcinoma, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus, and in particular, hematopoietic cancers such as lymphoma, leukemia, and myeloma; immune disorders such as actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, and trauma; infections such as those caused by viral agents classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus, hepadnavirus, herpesvirus, flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, and togavirus; infections caused by bacterial agents classified as pneumococcus, staphylococcus, streptococcus, bacillus, corynebacterium, clostridium, meningococcus, gonococcus, listeria, moraxella, kingella, haemophilus, legionella, bordetella, gram-negative enterobacterium including shigella, salmonella, and campylobacter, pseudomonas, vibrio, brucella, francisella, yersinia, bartonella, norcardium, actinomyces, mycobacterium, spirochaetale, rickettsia, chlamydia, and mycoplasma; infections caused by fungal agents classified as aspergillus, blastomyces, dermatophytes, cryptococcus, coccidioides, malasezzia, histoplasma, and other fungal agents causing various mycoses; infections caused by parasites classified as plasmodium or malaria-causing, parasitic entamoeba, leishmania, trypanosoma, toxoplasma, pneumocystis carinii,

intestinal protozoa such as giardia, trichomonas, tissue nematodes such as trichinella, intestinal nematodes such as ascaris, lymphatic filarial nematodes, trematodes such as schistosoma, and cestrodes such as tapeworm; and vascular disorders such as arteriovenous fistula, atherosclerosis including atherosclerotic coronary artery disease, arteriosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis, phlebothrombosis, vascular tumors, complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass, cardiovascular disease, heart failure, heart disease, angina pectoris, myocardial infarction, calcific aortic valve stenosis and other aortic valve disorders, endocarditis, carcinoid heart disease, and complications of cardiac transplantation.

In another embodiment, a vector capable of expressing BMDSP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of BMDSP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified BMDSP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of BMDSP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of BMDSP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of BMDSP including, but not limited to, those listed above.

In a further embodiment, an antagonist of BMDSP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of BMDSP. Examples of such disorders include, but are not limited to, those cancers, immune disorders, infections, and vascular disorders described above. In one aspect, an antibody which specifically binds BMDSP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express BMDSP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding BMDSP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of BMDSP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to

achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of BMDSP may be produced using methods which are generally known in the art. In particular, purified BMDSP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind BMDSP. Antibodies to BMDSP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with BMDSP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to BMDSP have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of BMDSP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to BMDSP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and

Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce BMDSP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for BMDSP may also be generated. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between BMDSP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering BMDSP epitopes is preferred, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for BMDSP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of BMDSP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple BMDSP epitopes, represents the average affinity, or avidity, of the antibodies for BMDSP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular BMDSP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the BMDSP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7

L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of BMDSP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of BMDSP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding BMDSP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding BMDSP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding BMDSP. Thus, complementary molecules or fragments may be used to modulate BMDSP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding BMDSP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding BMDSP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding BMDSP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding BMDSP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or

regulatory regions of the gene encoding BMDSP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding BMDSP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding BMDSP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as

inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of BMDSP, antibodies to BMDSP, and mimetics, agonists, antagonists, or inhibitors of BMDSP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by

the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a

manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of BMDSP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example BMDSP or fragments thereof, antibodies of BMDSP, and agonists, antagonists or inhibitors of BMDSP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the

subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind BMDSP may be used for the diagnosis of disorders characterized by expression of BMDSP, or in assays to monitor patients being treated with BMDSP or agonists, antagonists, or inhibitors of BMDSP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for BMDSP include methods which utilize the antibody and a label to detect BMDSP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring BMDSP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of BMDSP expression. Normal or standard values for BMDSP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to BMDSP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, preferably by photometric means. Quantities of BMDSP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding BMDSP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide

sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of BMDSP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of BMDSP, and to monitor regulation of BMDSP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding BMDSP or closely related molecules may be used to identify nucleic acid sequences which encode BMDSP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding BMDSP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the BMDSP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:3-4 or from genomic sequences including promoters, enhancers, and introns of the BMDSP gene.

Means for producing specific hybridization probes for DNAs encoding BMDSP include the cloning of polynucleotide sequences encoding BMDSP or BMDSP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding BMDSP may be used for the diagnosis of disorders associated with expression of BMDSP. Examples of such disorders include, but are not limited to, cancers such as adenocarcinoma, melanoma, sarcoma, teratocarcinoma, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus, and in particular, hematopoietic cancers such as lymphoma, leukemia, and myeloma; immune disorders such as actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis,

autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, and trauma; infections such as those caused by viral agents classified as adenovirus, arenavirus, bunyavirus, calicivirus, coronavirus, filovirus, hepadnavirus, herpesvirus, flavivirus, orthomyxovirus, parvovirus, papovavirus, paramyxovirus, picornavirus, poxvirus, reovirus, retrovirus, rhabdovirus, and togavirus; infections caused by bacterial agents classified as pneumococcus, staphylococcus, streptococcus, bacillus, corynebacterium, clostridium, meningococcus, gonococcus, listeria, moraxella, kingella, haemophilus, legionella, bordetella, gram-negative enterobacterium including shigella, salmonella, and campylobacter, pseudomonas, vibrio, brucella, francisella, yersinia, bartonella, norcardium, actinomyces, mycobacterium, spirochaetale, rickettsia, chlamydia, and mycoplasma; infections caused by fungal agents classified as aspergillus, blastomyces, dermatophytes, cryptococcus, coccidioides, malasezzia, histoplasma, and other fungal agents causing various mycoses; infections caused by parasites classified as plasmodium or malaria-causing, parasitic entamoeba, leishmania, trypanosoma, toxoplasma, pneumocystis carinii, intestinal protozoa such as giardia, trichomonas, tissue nematodes such as trichinella, intestinal nematodes such as ascaris, lymphatic filarial nematodes, trematodes such as schistosoma, and cestrodes such as tapeworm; and vascular disorders such as arteriovenous fistula, atherosclerosis including atherosclerotic coronary artery disease, arteriosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis, phlebothrombosis, vascular tumors, complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass, cardiovascular disease, heart failure, heart disease, angina pectoris, myocardial infarction, calcific aortic valve stenosis and other aortic valve disorders, endocarditis, carcinoid heart disease, and complications of cardiac transplantation. The polynucleotide sequences encoding BMDSP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and

multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered BMDSP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding BMDSP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding BMDSP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding BMDSP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of BMDSP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding BMDSP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding

BMDSP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding BMDSP, or a fragment of a polynucleotide complementary to the polynucleotide encoding BMDSP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of BMDSP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding BMDSP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in

Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding BMDSP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, BMDSP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between BMDSP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with BMDSP, or fragments thereof, and washed. Bound BMDSP is then detected by methods well known in the art. Purified BMDSP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding BMDSP specifically compete with a test compound for binding BMDSP. In this manner, antibodies can be used to detect the presence of any peptide

which shares one or more antigenic determinants with BMDSP.

In additional embodiments, the nucleotide sequences which encode BMDSP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. [Attorney Docket No. PF-0609 P, filed October 2, 1998], are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

BMARNOT02

The BMARNOT02 library was constructed using RNA purchased from Clontech. The RNA was isolated from the bone marrow of 24 male and female Caucasian donors, 16 to 70 years old. The cDNA library was custom constructed by Stratagene using this RNA. cDNA synthesis was primed using oligo d(T) and random hexamers, and the cDNA library was cloned using the UNIZAP vector system (Stratagene).

PROSNOT18

The PROSNOT18 library was constructed using RNA isolated from diseased prostate tissue removed from a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, and gastrostomy. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated grade 3 transitional cell carcinoma. Patient history included angina and emphysema. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.

Frozen tissue was homogenized and lysed in guanidinium isothiocyanate solution using a Polytron PT-3000 homogenizer (Brinkmann Instruments, Westbury NY). The lysate was centrifuged over a CsCl cushion to isolate RNA. The RNA was extracted with acid phenol, precipitated with sodium acetate and ethanol, resuspended in RNase-free water, and treated with

DNase. The RNA was re-extracted with acid phenol and reprecipitated as described above. Poly(A+) RNA was isolated using the OLIGOTEX mRNA purification kit (QIAGEN, Chatsworth CA).

Poly(A+) RNA was used for cDNA synthesis and construction of the cDNA library according to the recommended protocols in the SUPERSCRIPT plasmid system (Life Technologies). The cDNAs were fractionated on a SEPHAROSE CL4B column (Amersham Pharmacia Biotech), and those cDNAs exceeding 400 bp were ligated into pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into DH5α competent cells (Life Technologies).

II. Isolation of cDNA Clones

BMARNOT02

cDNA clones were recovered by in vivo excision as single-stranded PBLUESCRIPT phagemids (Stratagene). These phagemids were used to reinfect SOLR host cells (Stratagene) from which double-stranded recombinant phagemids were purified using either the QIAWELL-8 plasmid purification system (QIAGEN) or the MINIPREP plasmid purification kit (Advanced Genetic Technologies Corp., Gaithersburg, MD).

PROSNOT18

Plasmid DNA was released from the cells and purified using the R.E.A.L. Prep 96 plasmid kit (QIAGEN). The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Life Technologies) with carbenicillin at 25 mg/l and glycerol at 0.4%; 2) after the cultures were incubated for 19 hours, the cells were lysed with 0.3 ml of lysis buffer; and 3) following isopropanol precipitation, the plasmid DNA pellets were each resuspended in 0.1 ml of distilled water. The DNA samples were stored at 4°C.

III. Sequencing and Analysis

The cDNAs were prepared for sequencing using the ABI CATALYST 800 (Perkin-Elmer) or the HYDRA microdispenser (Robbins Scientific) or MICROLAB 2200 (Hamilton) systems in combination with the PTC-200 thermal cyclers (MJ Research). The cDNAs were sequenced using the ABI PRISM 373 or 377 sequencing systems (Perkin-Elmer) and standard ABI protocols, base calling software, and kits. In one alternative, cDNAs were sequenced using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics). In another alternative, the cDNAs were amplified and sequenced using the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). In yet another alternative, cDNAs were sequenced using solutions and dyes from Amersham Pharmacia Biotech. Reading frames for the ESTs were determined using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA

sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA, extension, and shotgun sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 1 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 1 shows the tools, programs, and algorithms used, the second column provides brief descriptions, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering) and LASERGENE software (DNASTAR).

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases, such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, PFAM, and Prosite.

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:3-4. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or

related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\%\text{ sequence identity} \times \%\text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding BMDSP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation/trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in The Invention section.

V. Extension of BMDSP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:3-4 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+},

$(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulphoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO.3-4 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:3-4 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized and compared using autoradiography.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the

present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the BMDSP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring BMDSP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of BMDSP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the BMDSP-encoding transcript.

IX. Expression of BMDSP

Expression and purification of BMDSP is achieved using bacterial or virus-based expression systems. For expression of BMDSP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express BMDSP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of BMDSP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding BMDSP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, BMDSP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from BMDSP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified BMDSP obtained by these methods can be used directly in the following activity assay.

X. Demonstration of BMDSP Activity

An assay for BMDSP-1 activity measures the precipitation of antigen from serum using the quantitative precipitin reaction (Golub, E.S. et al. (1987) Immunology: A Synthesis, Sinauer Associates, Sunderland MA, pp. 113-115). BMDSP-1 is isotopically labeled using methods known in the art. Various serum concentrations are added to constant amounts of labeled BMDSP-1. BMDSP-1/antigen complexes precipitate out of solution and are collected by centrifugation. The amount of precipitable BMDSP-1/antigen complex is proportional to the amount of radioisotope detected in the precipitate. The amount of precipitable BMDSP-1/antigen complex is plotted against the serum concentration. For various serum concentrations, a characteristic precipitin curve is obtained, in which the amount of precipitable BMDSP-1/antigen complex initially increases proportionately with increasing serum concentration, peaks at the equivalence point, and then decreases proportionately with further increases in serum concentration. Thus, the amount of precipitable BMDSP-1/antigen complex is a measure of BMDSP-1 activity and is characterized by sensitivity to both limiting and excess quantities of antigen.

An assay for BMDSP-2 activity measures the ability of BMDSP-2 to complex with proteins from bone marrow stromal cells. BMDSP-2, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton, A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Stromal cell-specific proteins previously arrayed in the wells of a microtiter plate are incubated with the labeled BMDSP-2 and washed. The amount of labeled BMDSP-2 in each well is quantified and is proportional to the amount of BMDSP-2/stromal cell protein complex. Data obtained using different concentrations of labeled BMDSP-2 are used to

calculate the affinity of BMDSP-2 for stromal cell proteins.

XI. Functional Assays

BMDSP function is assessed by expressing the sequences encoding BMDSP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP, and to evaluate properties, for example, their apoptotic state. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of BMDSP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding BMDSP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding BMDSP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of BMDSP Specific Antibodies

BMDSP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the BMDSP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A Peptide Synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring BMDSP Using Specific Antibodies

Naturally occurring or recombinant BMDSP is substantially purified by immunoaffinity chromatography using antibodies specific for BMDSP. An immunoaffinity column is constructed by covalently coupling anti-BMDSP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing BMDSP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of BMDSP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/BMDSP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and BMDSP is collected.

XIV. Identification of Molecules Which Interact with BMDSP

BMDSP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent (Bolton, supra). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled BMDSP, washed, and any wells with labeled BMDSP complex are assayed. Data obtained using different concentrations of BMDSP are used to calculate values for the number, affinity, and association of BMDSP with the candidate molecules.

Various modifications and variations of the described methods and systems of the

invention will be apparent to those skilled in the art without depa.ing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS and PRINTS databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and Probability value= 1.0E-3 or less
PFAM	A Hidden Markov Models-based application useful for protein family search.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits, depending on individual protein families

Table 1 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Score= 4.0 or greater
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-2, and fragments thereof.

2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:3-4, and fragments thereof.

10. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in

conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of BMDSP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of BMDSP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/47, G01N 33/50, C12Q 1/68, C07K 16/18, A61K 38/17

A2

(11) International Publication Number: WO 00/20588

(43) International Publication Date: 13 April 2000 (13.04.00)

(21) International Application Number: PCT/US99/22908

(22) International Filing Date: 1 October 1999 (01.10.99)

(30) Priority Data:

09/165,621	2 October 1998 (02.10.98)	US
Not furnished	2 October 1998 (02.10.98)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	09/165,621 (CIP)
Filed on	2 October 1998 (02.10.98)
US	Not furnished (CIP)
Filed on	2 October 1998 (02.10.98)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94087 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: BONE MARROW-DERIVED SERUM PROTEINS

```
1    MPAK------TPIYL-KAANNKKGKKFKLR  1859631
1    MPGPQGGRGAATMSLGKLSPVGWVSSSQGK  GI 338033

24   DILSPDMISPPLGDFRHTIHIGKEGQHDVF  1859631
31   RRLTADMISHPLGDFRHTMHVGRGG--DVF  GI 338033

54   GDISFLQGNYELLPGNQEKAHLGQFPGHNE  1859631
59   GDTSFL-SNHGGSSGSTHRSPRSFLAKKLQ  GI 338033

84   FFRANSTSDSVFTETPSPVLKNAISLPTIG  1859631
88   LVRRVGAPPRRMASPPAPSPAPPAISPIII-  GI 338033

114  GSQALMLPLLSPVTFNSKQESFGPAKLPRL  1859631
117  -KNAISLPQLNQAAY----DSLVVGKLS-F  GI 338033
```

(57) Abstract

The invention provides human bone marrow-derived serum proteins (BMDSP) and polynucleotides which identify and encode BMDSP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating or preventing disorders associated with expression of BMDSP.

```
                11                  20                  29                  38                  47                  56
5' TCGAG CCG ATT CGG CTC GAG CGG CTC GAG CTC AGT TAG GAC CCA GAG GGA ACC ATG
                                                                          M

                65                  74                  83                  92                 101                 110
    GAA GCC CCA GCT CAG CTT CTC TTC CTC CTG CTA CTC TGG CTC CCA GAT ACC ACC
    E   A   P   A   Q   L   L   F   L   L   L   L   W   L   P   D   T   T

               119                 128                 137                 146                 155                 164
    GGA GAA ATT GTG TTG ACA CAG TCT CCA GCC ACC CTG TCT TTG TCT CCA GGG GAA
    G   E   I   V   L   T   Q   S   P   A   T   L   S   L   S   P   G   E

               173                 182                 191                 200                 209                 218
    AGA GCC ACC CTC TCC TGC AGG GCC AGT CAG AGT GTT AGC AGC TAC TTA GCC TGG
    R   A   T   L   S   C   R   A   S   Q   S   V   S   S   Y   L   A   W

               227                 236                 245                 254                 263                 272
    TAC CAA CAG AAA CCT GGC CAG GCT CCC AGG CTC CTC ATC TAT GAT GCA TCC AAC
    Y   Q   Q   K   P   G   Q   A   P   R   L   L   I   Y   D   A   S   N

               281                 290                 299                 308                 317                 326
    AGG GCC ACT GGC ATC CCA CCC AGG TTC AGT GGC AGT GGG TCT GGG ACA GAC TTC
    R   A   T   G   I   P   P   R   F   S   G   S   G   S   G   T   D   F

               335                 344                 353                 362                 371                 380
    ACT CTC ACC ATC AGC AGA CTG GAG CCC GAA GAT GTG GCA CTT TAT TAC TGT CAG
    T   L   T   I   S   R   L   E   P   E   D   V   A   L   Y   Y   C   Q
```

FIGURE 1A

```
                389        398        407        416        425        434
CAA TAT TTT ACT ACT CCG TAC ACT TTT GGC CAG GGG ACC AGG CTG GAG ATC AAA
Q   Y   F   T   T   P   Y   T   F   G   Q   G   T   R   L   E   I   K

                443        452        461        470        479        488
CGA ACT GTG GCT GCA CCA TCT GTC TTC ATC TTC CCG CCA TCT GAT GAG CAG TTG
R   T   V   A   A   P   S   V   F   I   F   P   P   S   D   E   Q   L

                497        506        515        524        533        542
AAA TCT GGA ACT GCC TCT GTT GTG TGC CTG CTG AAT AAC TTC TAT CCC AGA GAG
K   S   G   T   A   S   V   V   C   L   L   N   N   F   Y   P   R   E

                551        560        569        578        587        596
GCC AAA GTA CAG TGG AAG GTG GAT AAC GCC CTC CAA TCG GGT AAC TCC CAG GAG
A   K   V   Q   W   K   V   D   N   A   L   Q   S   G   N   S   Q   E

                605        614        623        632        641        650
AGT GTC ACA GAG CAG GAC AGC AAG GAC AGC ACC TAC AGC CTC AGC AGC ACC CTG
S   V   T   E   Q   D   S   K   D   S   T   Y   S   L   S   S   T   L

                659        668        677        686        695        704
ACG CTG AGC AAA GCA GAC TAC GAG AAA CAC AAA GTC TAC GCC TGC GAA GTC ACC
T   L   S   K   A   D   Y   E   K   H   K   V   Y   A   C   E   V   T

                713        722        731        740        749        758
CAT CAG GGC CTG AGC TCG CCC GTC ACA AAG AGC TTC AAC AGG GGA GAG TGT TAG
H   Q   G   L   S   S   P   V   T   K   S   F   N   R   G   E   C
```

FIGURE 1B

```
        767         776         785         794         803         812
AGG GAG AAG TGC CCC CAC CTG CTC CTC AGT TCC AGC CTG ACC CCC TCC CAT CCT

        821         830         839         848         857         866
TTG GCC TCT GAC CCT TTT TCC ACA GGG GAC CTA CCC CTA TTG CGG TCC TCC AGC

        875         884         893         902         911         920
TCA TCT TTC ACC TCA CCC CCC TCC TCC TCC TTG GCT TTA ATT ATG CTA ATG TTG

        929         938         947         956
GAG GAG AAT GAA TAA ATA AAG TGA ATC TTT GCA AAA AAA AAA  3'
```

FIGURE 1C

```
                10                  19                  28                  37                  46                  55
5' CGGG GCT AGC CCG GAG ACC CGG CCA CCG GCC TGG GGC GCC TTC ACG CCG TCT CGG

                64                  73                  82                  91                 100                 109
AGC GGA TAA TGC GGT GAG CAG GCA CCA CGC CGG CAG ACT CGG CTG GAT CTG CGC

               118                 127                 136                 145                 154                 163
ACA GCG GCA GGG ATT GCG TGC GCC CGC GGG AGG CCC GGG GCA GCG GCT GGG ATC

               172                 181                 190                 199                 208                 217
CTC AGC GGC CGG TTT GTC CTG GTT GTG GTC AAG ACT GGA TGA TGT AAC TGG

               226                 235                 244                 253                 262                 271
CTC TCT AGG AAG CCT CAC TTG GCC GTA ACC TCA GGA AGG TTC TCT TTG ACC CCA

               280                 289                 298                 307                 316                 325
TCT CAT TTC GAA GCC ACT TCT GAA GCC ACT TGA GAA AAA TGA TGT GAC AGT TCC

               334                 343                 352                 361                 370                 379
TAT CAA AAA GGA TTC AGA AAC ATA TAC CAT CTG TGA AGA AAG TGG CCC TTT CTC

               388                 397                 406                 415                 424                 433
CCG CTT GCA AAA TAG ACA TTC TCA AAT TCC AAA ATG CCA GCC AAG ACC CCA ATT
                                                        M   P   A   K   T   P   I
```

FIGURE 2A

```
                442                 451                 460                 469                 478                 487
TAC CTG AAA GCA GCC AAT AAC AAG AAA GGA AAG AAA TTT AAA CTG AGG GAC ATT
Y   L   K   A   A   N   N   K   K   G   K   K   F   K   L   R   D   I

                496                 505                 514                 523                 532                 541
CTG TCT CCT GAT ATG ATC AGT CCC CCG CTT GGA GAC TTT CGC CAC ACC ATC CAC
L   S   P   D   M   I   S   P   P   L   G   D   F   R   H   T   I   H

                550                 559                 568                 577                 586                 595
ATT GGC AAA GAG GGC CAG CAC GAT GTC TTT GGA GAT ATT TCC TTT CTT CAA GGG
I   G   K   E   G   Q   H   D   V   F   G   D   I   S   F   L   Q   G

                604                 613                 622                 631                 640                 649
AAC TAC GAG CTT TTA CCT GGA AAC CAG GAG AAA GCA CAC CTG GGC CAG TTC CCT
N   Y   E   L   L   P   G   N   Q   E   K   A   H   L   G   Q   F   P

                658                 667                 676                 685                 694                 703
GGG CAT AAT GAG TTC TTC CGG GCC AAC AGC ACC TCG GAC TCT GTG TTC ACA GAA
G   H   N   E   F   F   R   A   N   S   T   S   D   S   V   F   T   E

                712                 721                 730                 739                 748                 757
ACG CCC TCC CCG GTG CTC AAA AAT GCC ATC TCC CTC CCG ACC ATT GGA GGA TCC
T   P   S   P   V   L   K   N   A   I   S   L   P   T   I   G   G   S

                766                 775                 784                 793                 802                 811
CAA GCT CTC ATG TTG CCC TTA TTG TCA CCA GTG ACA TTT AAT TCC AAA CAG GAG
Q   A   L   M   L   P   L   L   S   P   V   T   F   N   S   K   Q   E
```

FIGURE 2B

```
      820        829        838        847        856        865
TCC TTC GGG CCA GCA AAG CTG CCC AGG CTT AGC TGC GAG CCC GTC ATG GAG GAA
S   F   G   P   A   K   L   P   R   L   S   C   E   P   V   M   E   E

      874        883        892        901        910        919
AAA GCT CAG GAG AAA AGC AGT CTG TTG GAG AAT GGG ACA GTC CAC CAG GGA GAC
K   A   Q   E   K   S   S   L   L   E   N   G   T   V   H   Q   G   D

      928        937        946        955        964        973
ACC TCG TGG GGC TCC AGC GGT TCT GCA TCT CAG TCC AGC CAA GGC AGA GAC AGC
T   S   W   G   S   S   G   S   A   S   Q   S   S   Q   G   R   D   S

      982        991       1000       1009       1018       1027
CAC TCC TCC AGC CTG TCC GAA CAG TAC CCC GAC TGG CCA GCC GAG GAC ATG TTT
H   S   S   S   L   S   E   Q   Y   P   D   W   P   A   E   D   M   F

     1036       1045       1054       1063       1072       1081
GAC CAT CCC ACC CCA TGC GAG CTC ATC AAG GGA AAG ACT AAG TCA GAG GAG TCC
D   H   P   T   P   C   E   L   I   K   G   K   T   K   S   E   E   S

     1090       1099       1108       1117       1126       1135
CTC TCT GAC CTT ACA GGT TCC CTC CTC TCC CTG CAG CTT GAT CTT GGG CCC TCA
L   S   D   L   T   G   S   L   L   S   L   Q   L   D   L   G   P   S

     1144       1153       1162       1171       1180       1189
CTT TTG GAT GAG GTG CTG AAT GTA ATG GAT AAA AAT AAG TAA CAA GAT GCC AAC
L   L   D   E   V   L   N   V   M   D   K   N   K
```

FIGURE 2C

```
1198                1207                1216                1225                1234                1243
TTT TTT CCT TTG GGG TAA AAG GTA CAA AAA CAA ACT AAC CAC AGT TGA AGA GAA

1252                1261                1270                1279                1288                1297
GGG CTT CCG GAG CTG TAT TTG CAG TTT TGT GTT GGG TTT TCT AAA ATA ATA TTC

1306                1315                1324                1333                1342                1351
TTA CAA AGT ATT TTT TTA CCT GTT ATG CCC TGT TTG CAA AAA CAA TTT AGA AAA

1360                1369                1378                1387                1396                1405
AAA CAA CAA AGC AAA ACC TAT CTT GGC AAA AAA AGG AAG TGA GTC AGA GCC CAT

1414                1423                1432                1441                1450                1459
TTT CAG GAG GCA TTG GTG ATG TTC GGC TCA CAT ATT GTT TGC AGA CAC ACA AGA

1468                1477                1486                1495                1504                1513
AAT CTG GCT TGG CCA GGA TTG GCA CTA GCT ATG AAG GGC TGA GCG AGT CAC ATT

1522                1531                1540                1549                1558                1567
AAG GAA CTT CAC GGA ACT TTA TAG CAC TCC GAC ATT TTC TGA GCA AGA GGA AGT

1576                1585                1594                1603                1612                1621
CAA AAT TTA TTT AAC ACC TAA GCC TTT TTG TAG ACT CTT TTC TAT ATA TTG CTT

1630                1639                1648                1657                1666                1675
AGG CTC ACC ATA GCG AAT TCT CCA GTG TTA AAA CTT TTC TGT TTT CAC ATT TGA
```

FIGURE 2D

```
1684                1693                1702                1711                1720                1729
ACT TTA TGG GTT TTG GGG ATT TTC TTG TAG TTC TTA TAT ATC CCT ATA TAT TAT

1738                1747                1756                1765                1774                1783
ATC TAT ATT GCA AAA TTT TGA CTG TCA GCT ACA TGT TGG TAA GAC ACA GGC AAA

1792                1801                1810                1819                1828                1837
GTA TTA CTG TAA CTA AGT TAT TTT TAA AGT TAA AAT ATA TTT TTA CGT GCC TTT

1846                1855                1864                1873                1882                1891
GGC TTT TTA TTG CAG AGT CTA CAT TTT ATA GAT TCT ACA TCA GAT GTT GTC ACT

1900                1909                1918                1927                1936                1945
TAT TTC CAT TGG GAT TCC ATT GTA AGC TGT GTA TGT GCG TGT TTG GAA AAG TGT

1954                1963                1972                1981                1990                1999
ATT CAT ACT TAG TTT TTT CTT CAT CTG TTA TCA TAC TTT TAA CAG CAA CCA

2008                2017                2026                2035                2044                2053
ATA ACG GAT TGT AAA GTG TAA AGG CAC AGG TTA CTC ATG ATG CTT CTG CAG AGA

2062                2071                2080                2089                2098                2107
CTG TGG GCT ACA CCA CAT ATG TTA TTT GGA AAT ATA GGT ATT TTA GTA CAG TAC

2116                2125                2134                2143                2152                2161
ATA CTT GCA TTA CAT AGG TAC TTC AAG CAA CAC AAT AAA AAG TAA ATG ATA AAG
```

FIGURE 2E

```
      2170        2179        2188        2197        2206        2215
TGA ACT TGC TTG TTT ATA GTA ATA AAC AAG ACC ATA AGA GAA TAA GTA TAG CTA

      2224        2233        2242        2251        2260        2269
GAG AAA TTG CTT CTC TGA AAT GTA CAT GAG CCC TTA AGG TAA GAG ATG ATT TCC

      2278        2287        2296        2305        2314        2323
ATC TAC TCT CAT TTT GAT TAC TTC CTT ATG GTT TGA GAG GCT AGA AAC TGA GCC

      2332        2341        2350        2359        2368        2377
TCT CTA CTT TTG GAA AAA TGA ACA TGT GAG GTC AGA TTT TTT TTT TTT TTT TTA

      2386        2395        2404        2413        2422        2431
AGT CAG CAC TGA TGC CAC CCT CTC AGT GGT CAT TTC TGA GCA TCT TCC TGA CTT

      2440        2449        2458        2467        2476        2485
GAA CAC CTT CTA CAG CAA ACT CTT GCA AGT CCA GTT TCA TCC CTG TAA GGC AAA

      2494        2503        2512        2521        2530        2539
TGT CTT TTC ACG CAG AAA GTG CCA TAT AGA CGA GAT AAA GGC AGC TAN AAC GAG


GGC AGT A 3'
```

FIGURE 2F


1 MPAK - - - - - - TPIYL - KAANNKKGKKFKLR 1859631
1 MP GPQGGRGAATMSL G KLSPVGWVSSSQGGK GI 338033
24 DILSPDMISPPLGDFRHTIHIGKEGQHDVF 1859631
31 RRL TADMIS HPLGDFRHTM HV GRGG - - DVF GI 338033
54 GDISFLQGNYELLPGNQEKAHLGQFPGHNE 1859631
59 GD TSFL - S NHGGSSG STHRSPRSFLAKKLQ GI 338033
84 FFRANSTSDSVFTETPSPVLKNAISLPTIG 1859631
88 LV R RVGAPPRRMASP PA PSPAPPAISP I I - GI 338033
114 GSQALMLPLLSPVTFNSKQESFGPAKLPRL 1859631
117 - K N A I S L P Q L N Q A A Y - - - - D S L V V G K L S - F GI 338033

FIGURE 3A

```
144 SCEPVMEEKAQEKSSLLENGTVHQGDTSWG 1859631
141 DSSPTSSTDGHSSYGLL-DSGFCTISRLPR- GI 338033

174 SSGSASQSSQGRDSHSSSLSEQYPDWPAED 1859631
169 ----SEKPHDRDRDGS----FPSEP---- GI 338033

204 MFDHPTPCELIKGKTKSEESLSDLTGSLLS 1859631
186 ----------GLRRSDSLLS----------FR GI 338033

234 LQLDLGPSLLDEVLNVMDKNK 1859631
198 LDLDLGPSLLSELLGVMSLPE GI 338033
```

FIGURE 3B

Docket No.: PF-0609 USN

DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

BONE MARROW-DERIVED SERUM PROTEINS

the specification of which:

/___/ is attached hereto.

/___/ was filed on _____ as application Serial No. _____ and if this box contains an X /___/, was amended on _____.

/ X / was filed as Patent Cooperation Treaty international application No. PCT/US99/22908 on October 1, 1999, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Article 19 on _____ 2001, and if this box contains an X /_/, was amended on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

Docket No.: PF-0609 USN

Country	Number	Filing Date	Priority Claimed
_____	_____	_____	/_/ Yes /_/ No
_____	_____	_____	/_/ Yes /_/ No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
60/155,264	October 2, 1998	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)

I hereby appoint the following:

Lucy J. Billings	Reg. No. 36,749
Michael C. Cerrone	Reg. No. 39,132
Diana Hamlet-Cox	Reg. No. 33,302
Richard C. Ekstrom	Reg. No. 37,027
Barrie D. Greene	Reg. No. 46,740
Matthew R. Kaser	Reg. No. 44,817
Lynn E. Murry	Reg. No. 42,918
Shirley A. Recipon	Reg. No. 47,016
Susan K. Sather	Reg. No. 44,316
Michelle M. Stempien	Reg. No. 41,327
David G. Streeter	Reg. No. 43,168
Stephen Todd	Reg. No. 47,139
Christopher Turner	Reg. No. 45,167
P. Ben Wang	Reg. No. 41,420

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

Docket No.: PF-0609 USN

LEGAL DEPARTMENT
INCYTE GENOMICS, INC.
3160 PORTER DRIVE, PALO ALTO, CA 94304

TEL: 650-855-0555 FAX: 650-849-8886 or 650-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

1-00

Sole Inventor or First Joint Inventor:

Full name: Y. TOM TANG

Signature: Y. Tom Tang

Date: Febmay 27, 2001

Citizenship ~~People's Republic of China~~ USA Y.T.T.

Residence: San Jose, California CA. 02/27/2001

P.O. Address: 4230 Ranwick Court

San Jose, California 95118

2-00

Second Joint Inventor:

Full name: NEIL C. CORLEY

Signature: Neil C Corley

Date: March 5, 2001

Citizenship United States of America

Residence: Castro Valley, California CA.

P.O. Address: 20426 Crow Creek Road

Castro Valley, California 94552

Docket No.: PF-0609 USN

3-00

Third Joint Inventor:

Full name: KARL J. GUEGLER

Signature:

Date: 02/02 , 2001

Citizenship Switzerland

Residence: Menlo Park, California CA.

P.O. Address: 1048 Oakland Avenue
Menlo Park, California 94025

4-00

Fourth Joint Inventor:

Full name: DYUNG AINA M. LU

Signature:

Date: March 22 , 2001

Citizenship United States of America

Residence: San Jose, California CA.

P.O. Address: 233 Coy Drive
San Jose, California 95123

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
TANG, Y. Tom
CORLEY, Neil C.
GUEGLER, Karl J.
LU, Aina Dyung M.

<120> BONE MARROW-DERIVED SERUM PROTEINS

<130> PF-0609 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/165,621; unassigned
<151> 1998-10-02; 1999-10-02

<160> 5

<170> PERL Program

<210> 1
<211> 234
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 135698CD1

<400> 1

```
Met Glu Ala Pro Ala Gln Leu Leu Phe Leu Leu Leu Leu Trp Leu
  1               5                  10                  15
Pro Asp Thr Thr Gly Glu Ile Val Leu Thr Gln Ser Pro Ala Thr
             20                  25                  30
Leu Ser Leu Ser Pro Gly Glu Arg Ala Thr Leu Ser Cys Arg Ala
             35                  40                  45
Ser Gln Ser Val Ser Ser Tyr Leu Ala Trp Tyr Gln Gln Lys Pro
             50                  55                  60
Gly Gln Ala Pro Arg Leu Leu Ile Tyr Asp Ala Ser Asn Arg Ala
             65                  70                  75
Thr Gly Ile Pro Pro Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp
             80                  85                  90
Phe Thr Leu Thr Ile Ser Arg Leu Glu Pro Glu Asp Val Ala Leu
             95                 100                 105
Tyr Tyr Cys Gln Gln Tyr Phe Thr Thr Pro Tyr Thr Phe Gly Gln
            110                 115                 120
Gly Thr Arg Leu Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val
            125                 130                 135
Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala
            140                 145                 150
Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys
            155                 160                 165
Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
            170                 175                 180
```

```
Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu
                185                 190                 195
Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys
                200                 205                 210
Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val
                215                 220                 225
Thr Lys Ser Phe Asn Arg Gly Glu Cys
                230
```

<210> 2
<211> 254
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1859631CD1

<400> 2

```
Met Pro Ala Lys Thr Pro Ile Tyr Leu Lys Ala Ala Asn Asn Lys
  1               5                  10                  15
Lys Gly Lys Lys Phe Lys Leu Arg Asp Ile Leu Ser Pro Asp Met
                20                  25                  30
Ile Ser Pro Pro Leu Gly Asp Phe Arg His Thr Ile His Ile Gly
                35                  40                  45
Lys Glu Gly Gln His Asp Val Phe Gly Asp Ile Ser Phe Leu Gln
                50                  55                  60
Gly Asn Tyr Glu Leu Leu Pro Gly Asn Gln Glu Lys Ala His Leu
                65                  70                  75
Gly Gln Phe Pro Gly His Asn Glu Phe Phe Arg Ala Asn Ser Thr
                80                  85                  90
Ser Asp Ser Val Phe Thr Glu Thr Pro Ser Pro Val Leu Lys Asn
                95                 100                 105
Ala Ile Ser Leu Pro Thr Ile Gly Gly Ser Gln Ala Leu Met Leu
                110                 115                 120
Pro Leu Leu Ser Pro Val Thr Phe Asn Ser Lys Gln Glu Ser Phe
                125                 130                 135
Gly Pro Ala Lys Leu Pro Arg Leu Ser Cys Glu Pro Val Met Glu
                140                 145                 150
Glu Lys Ala Gln Glu Lys Ser Ser Leu Leu Glu Asn Gly Thr Val
                155                 160                 165
His Gln Gly Asp Thr Ser Trp Gly Ser Ser Gly Ser Ala Ser Gln
                170                 175                 180
Ser Ser Gln Gly Arg Asp Ser His Ser Ser Ser Leu Ser Glu Gln
                185                 190                 195
Tyr Pro Asp Trp Pro Ala Glu Asp Met Phe Asp His Pro Thr Pro
                200                 205                 210
Cys Glu Leu Ile Lys Gly Lys Thr Lys Ser Glu Glu Ser Leu Ser
                215                 220                 225
Asp Leu Thr Gly Ser Leu Leu Ser Leu Gln Leu Asp Leu Gly Pro
                230                 235                 240
Ser Leu Leu Asp Glu Val Leu Asn Val Met Asp Lys Asn Lys
                245                 250
```

<210> 3
<211> 962
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 135698CB1

<400> 3
tcgagccgat tcggctcgag cggctcgagc tcagttagga cccagaggga accatggaag 60
ccccagctca gcttctcttc ctcctgctac tctggctccc agataccacc ggagaaattg 120
tgttgacaca gtctccagcc accctgtctt tgtctccagg ggaaagagcc accctctcct 180
gcagggccag tcagagtgtt agcagctact tagcctggta ccaacagaaa cctggccagg 240
ctcccaggct cctcatctat gatgcatcca acagggccac tggcatccca cccaggttca 300
gtggcagtgg gtctgggaca gacttcactc tcaccatcag cagactggag cccgaagatg 360
tggcacttta ttactgtcag caatatttta ctactccgta cacttttggc caggggacca 420
ggctggagat caaacgaact gtggctgcac catctgtctt catcttcccg ccatctgatg 480
agcagttgaa atctggaact gcctctgttg tgtgcctgct gaataacttc tatcccagag 540
aggccaaagt acagtggaag gtggataacg ccctccaatc gggtaactcc caggagagtg 600
tcacagagca ggacagcaag gacagcacct acagcctcag cagcaccctg acgctgagca 660
aagcagacta cgagaaacac aaagtctacg cctgcgaagt cacccatcag ggcctgagct 720
cgcccgtcac aaagagcttc aacaggggag agtgttagag ggagaagtgc ccccacctgc 780
tcctcagttc cagcctgacc ccctcccatc ctttggcctc tgacccttt tccacagggg 840
acctaccccct attgcggtcc tccagctcat ctttcacctc accccctcc tcctccttgg 900
ctttaattat gctaatgttg gaggagaatg aataaataaa gtgaatcttt gcaaaaaaaa 960
aa 962

<210> 4
<211> 2546
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 2533
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte ID No: 1859631CB1

<400> 4
cggggctagc ccggagaccc ggccaccggc ctggggcgcc ttcacgccgt ctcggagcgg 60
ataatgcggt gagcaggcac cacgccggca gactcggctg gatctgcgca cagcggcagg 120
gattgcgtgc gcccgcggga ggcccggggc agcggctggg atcctcagcg gcggccggtt 180
tgtcctggtt gtggtcaaga ctggatgatg taactggctc tctaggaagc ctcacttggc 240
cgtaacctca ggaaggttct ctttgacccc atctcatttc gaagccactt ctgaagccac 300
ttgagaaaaa tgatgtgaca gttcctatca aaaaggattc agaaacatat accatctgtg 360
aagaaagtgg ccctttctcc cgcttgcaaa atagacattc tcaaattcca aaatgccagc 420
caagacccca atttacctga aagcagccaa taacaagaaa ggaaagaaat ttaaactgag 480
ggacattctg tctcctgata tgatcagtcc cccgcttgga gactttcgcc acaccatcca 540
cattggcaaa gagggccagc acgatgtctt tggagatatt tcctttcttc aagggaacta 600
cgagctttta cctggaaacc aggagaaagc acacctgggc cagttccctg ggcataatga 660
gttcttccgg gccaacagca cctcggactc tgtgttcaca gaaacgccct ccccggtgct 720
caaaaatgcc atctccctcc cgaccattgg aggatcccaa gctctcatgt tgcccttatt 780
gtcaccagtg acatttaatt ccaaacagga gtccttcggg ccagcaaagc tgcccaggct 840

tagctgcgag cccgtcatgg aggaaaaagc tcaggagaaa agcagtctgt tggagaatgg 900
gacagtccac cagggagaca cctcgtgggg ctccagcggt tctgcatctc agtccagcca 960
aggcagagac agccactcct ccagcctgtc cgaacagtac cccgactggc cagccgagga 1020
catgtttgac catcccaccc catgcgagct catcaaggga aagactaagt cagaggagtc 1080
cctctctgac cttacaggtt ccctcctctc cctgcagctt gatcttgggc cctcactttt 1140
ggatgaggtg ctgaatgtaa tggataaaaa taagtaacaa gatgccaact tttttccttt 1200
ggggtaaaag gtacaaaaac aaactaacca cagttgaaga gaagggcttc cggagctgta 1260
tttgcagttt tgtgttgggt tttctaaaat aatattctta caaagtattt ttttacctgt 1320
tatgccctgt ttgcaaaaac aatttagaaa aaaacaacaa agcaaaacct atcttggcaa 1380
aaaaaggaag tgagtcagag cccattttca ggaggcattg gtgatgttcg gctcacatat 1440
tgtttgcaga cacacaagaa atctggcttg gccaggattg gcactagcta tgaagggctg 1500
agcgagtcac attaaggaac ttcacggaac tttatagcac tccgacattt tctgagcaag 1560
aggaagtcaa aatttattta acacctaagc ctttttgtag actcttttct atatattgct 1620
taggctcacc atagcgaatt ctccagtgtt aaaacttttc tgttttcaca tttgaacttt 1680
atgggttttg gggatttttct tgtagttctt atatatccct atatattata tctatattgc 1740
aaaatttttga ctgtcagcta catgttggta agacacaggc aaagtattac tgtaactaag 1800
ttatttttaa agttaaaata tatttttacg tgcctttggc tttttattgc agagtctaca 1860
ttttatagat tctacatcag atgttgtcac ttatttccat tgggattcca ttgtaagctg 1920
tgtatgtgcg tgtttggaaa agtgtattca tacttagttt ttttttcttc atctgttatc 1980
atacttttaa cagcaaccaa taacggattg taaagtgtaa aggcacaggt tactcatgat 2040
gcttctgcag agactgtggg ctacaccaca tatgttattt ggaaatatag gtattttagt 2100
acagtacata cttgcattac ataggtactt caagcaacac aataaaaagt aaatgataaa 2160
gtgaacttgc ttgtttatag taataaacaa gaccataaga gaataagtat agctagagaa 2220
attgcttctc tgaaatgtac atgagccctt aaggtaagag atgatttcca tctactctca 2280
ttttgattac ttccttatgg tttgagaggc tagaaactga gcctctctac ttttggaaaa 2340
atgaacatgt gaggtcagat tttttttttt ttttttaagt cagcactgat gccacctct 2400
cagtggtcat ttctgagcat cttcctgact tgaacacctt ctacagcaaa ctcttgcaag 2460
tccagtttca tccctgtaag gcaaatgtct tttcacgcag aaagtgccat atagacgaga 2520
taaaggcagc tanaacgagg gcagta 2546

<210> 5
<211> 391
<212> PRT
<213> Homo sapiens

<300>
<308> GenBank ID No: g338033

<400> 5

Met	Pro	Gly	Pro	Gln	Gly	Gly	Arg	Gly	Ala	Ala	Thr	Met	Ser	Leu	
1				5					10						15
Gly	Lys	Leu	Ser	Pro	Val	Gly	Trp	Val	Ser	Ser	Ser	Gln	Gly	Lys	
				20					25						30
Arg	Arg	Leu	Thr	Ala	Asp	Met	Ile	Ser	His	Pro	Leu	Gly	Asp	Phe	
				35					40						45
Arg	His	Thr	Met	His	Val	Gly	Arg	Gly	Gly	Asp	Val	Phe	Gly	Asp	
				50					55						60
Thr	Ser	Phe	Leu	Ser	Asn	His	Gly	Gly	Ser	Ser	Gly	Ser	Thr	His	
				65					70						75
Arg	Ser	Pro	Arg	Ser	Phe	Leu	Ala	Lys	Lys	Leu	Gln	Leu	Val	Arg	
				80					85						90
Arg	Val	Gly	Ala	Pro	Pro	Arg	Arg	Met	Ala	Ser	Pro	Pro	Ala	Pro	
				95					100						105
Ser	Pro	Ala	Pro	Pro	Ala	Ile	Ser	Pro	Ile	Ile	Lys	Asn	Ala	Ile	
				110					115						120

```
Ser Leu Pro Gln Leu Asn Gln Ala Ala Tyr Asp Ser Leu Val Val
                125                 130                 135
Gly Lys Leu Ser Phe Asp Ser Ser Pro Thr Ser Ser Thr Asp Gly
                140                 145                 150
His Ser Ser Tyr Gly Leu Asp Ser Gly Phe Cys Thr Ile Ser Arg
                155                 160                 165
Leu Pro Arg Ser Glu Lys Pro His Asp Arg Asp Arg Asp Gly Ser
                170                 175                 180
Phe Pro Ser Glu Pro Gly Leu Arg Arg Ser Asp Ser Leu Leu Ser
                185                 190                 195
Phe Arg Leu Asp Leu Asp Leu Gly Pro Ser Leu Leu Ser Glu Leu
                200                 205                 210
Leu Gly Val Met Ser Leu Pro Glu Ala Pro Ala Ala Glu Thr Pro
                215                 220                 225
Ala Pro Ala Ala Asn Pro Pro Ala Pro Thr Ala Asn Pro Thr Gly
                230                 235                 240
Pro Ala Ala Asn Pro Pro Ala Thr Thr Ala Asn Pro Pro Ala Pro
                245                 250                 255
Ala Ala Asn Pro Ser Ala Pro Ala Ala Thr Pro Thr Gly Pro Ala
                260                 265                 270
Ala Asn Pro Pro Ala Pro Ala Ala Ser Ser Thr Pro His Gly His
                275                 280                 285
Cys Pro Asn Gly Val Thr Ala Gly Leu Gly Pro Val Ala Glu Val
                290                 295                 300
Lys Ser Ser Pro Val Gly Gly Gly Pro Arg Gly Pro Ala Gly Pro
                305                 310                 315
Ala Leu Gly Arg His Trp Gly Ala Gly Trp Asp Gly Gly His His
                320                 325                 330
Tyr Pro Glu Met Asp Ala Arg Gln Glu Arg Val Glu Val Leu Pro
                335                 340                 345
Gln Ala Arg Ala Ser Trp Glu Ser Leu Asp Glu Glu Trp Arg Ala
                350                 355                 360
Pro Gln Ala Gly Ser Arg Thr Pro Val Pro Ser Thr Val Gln Ala
                365                 370                 375
Asn Thr Phe Glu Phe Ala Asp Ala Glu Glu Asp Asp Glu Val Lys
                380                 385                 390
Val
```